DOCUMENT RESUME

ED 251 845 CS 208 698

AUTHOR Seminoff, Nancy Wiseman; Elliott, Brownlee
TITLE Understanding Literature: Questions to Consider.
PUB DATE [81]
NOTE 24p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Anthologies; *Content Analysis; English Instruction; Grade 11; High Schools; Instructional Materials; Literary Genres; *Questioning Techniques; *Textbook Evaluation; Textbook Preparation; *Textbook Research; *United States Literature

ABSTRACT

Since the questions included in literature anthologies play a key role in the high school English class questioning process, a study examined the questions in four eleventh-grade American literature anthologies. The anthologies met the following criteria: textbooks (1) contained selections organized chronologically, (2) were major compilations in hard cover, (3) were edited by one or more major editors, (4) contained questions in the student edition, and (5) were recognized by their competitors as leaders in the field. Four fiction and four nonfiction selections were randomly selected from each anthology; 8 to 14 poems per anthology were selected. All of the questions that accompanied each selection were classified by levels of meaning (literal, interpretive, or applied), and rated on a five-point scale for clarity, objectivity, structure, and sequence. The analysis revealed that the textbook questions in general were weak in both clarity and structure. In addition, the texts were all weak in sequence. There was also a preponderance of interpretive questions. As was expected, there were significant differences in objectivity among texts and among genres. One text had a significantly lower rating on objectivity than the other three. Questions concerning poetry had a lower rating on objectivity than the genres of fiction and nonfiction. The findings led to the conclusion that textbook editors need to give greater priority to the writing of high quality textbook questions. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
[ ] Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED251845

UNDERSTANDING LITERATURE: QUESTIONS TO CONSIDER

by Nancy Wiseman Seminoff

and Brownlee Elliott

The Problem

At least since the dialogues of Socrates, teachers have believed that questions are an important part of the teaching-learning process. Today, for example, over one-fifth of the classroom time in typical high school English classes is devoted to teacher-dominated question and answer recitation (Squire and Applebee, 1960); and a significant portion of the remaining time is devoted to such activities as student questions and tests with questions.

Moreover, questions asked during recitations are not the only questions found in high school English classrooms. Textbook publishers for years now have routinely included questions with the reading selections in literature anthologies. These questions play an important role in the high school English teaching-learning process (Wease, 1976). Many teachers assign these questions, presumably as study guides for homework reading; some students use these questions as study guides for homework reading; and many teachers use these questions as models for their own recitation questions.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Nancy Seminoff

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS208698

But are anthology questions, which play such a key role in the high school English class questioning processes, well-designed? And if they are sometimes not well-designed, how might they be improved? These are the two primary questions addressed in this study.

Background

Content reading specialists (among others) have stressed the useful role that well-designed questions, including textbook questions, can play in helping students to learn (Wease, 1976). Herber and Nelson (1975) and Sanders (1974b) believe that well-designed questions can do more than help a student master a specific literature assignment; they believe that such questions can help a student master the process of understanding literature generally.

Herber (1978) describes a useful procedure for classifying questions according to their "levels of meaning," a qualitative distinction based on the various kinds of intellectual processes which questions can provoke. In Herber's system, understanding is viewed as a three level process: literal; interpretive; and applied.

At the literal level, readers determine what the author has explicitly stated. At the interpretive level, readers look for relationships among statements within the materials, and from these intrinsic relationships derive various meanings. The meanings may have been derived from the readers' past experience with the outside world; but that meaning is not (yet) applied back to the outside world. Finally, at the applied level, readers take the meanings they have derived at the literal and

interpretive levels and apply those meanings to other areas outside the printed text.

Herber's system of classification helps define the purposes of teaching questions. A student of Herber's (Sanders 1974a; 1974b) has in addition identified five characteristics of teaching questions; these characteristics may be related to the effectivness of teaching questions. And, in turn, a student of Sanders (Seminoff, 1980) has developed scales for measuring four of these five characteristics. The characteristics which Sanders identified are focus, clarity, objectivity, structure, and sequence; Seminoff developed scales and procedures for measuring all but focus (which for technical reasons did not lend itself as readily to measurement).

According to Sanders, questions which leave "no uncertainty as to what the student must do" have clarity; the phrasing of the question indicates specifically what kind of response will be acceptable to the teacher. Questions with clarity will contain some "task word" (list, describe, compare) which clearly indicates how the student should respond to the question.

Clear questions also provide parameters which guide the scope of the student's response. For example, the question

> How is the main point of the poem "Ozymandias" reflected in events which have happened?

lacks clarity when compared with the question

List three events from your history book which illustrate the main point of the poem "Ozymandias."

The word "list" is a task word in the second question, and the phrases "three events" and "from your history book" provide parameters to guide the student's response.

Questions without objectivity reveal the questioner's particular point of view. Word choice, for example, can reveal a questioner's bias (Riegle, 1975). In the question, "What do you think of the Kent State slaughter?" the use of the word "slaughter" pre-supposes a value assumption; hence, the question is not as objective as it could be.

Structure provides students with possible choices for responding to the task, rather than asking them to initiate their own responses. Thus, the question "What are some examples of Hamlet's ambiguous character?" lacks structure when compared with "From the list of actions below, chose those which seem to demonstrate Hamlet's 'shrewdness', and those which seem to demonstrate his 'insanity.'" One important type of assistance which contributes to structure is the provision of page, paragraph, or line numbers to direct students to specific information.

Sequence is the quality of questions presented in a logical (or psychological) order, "later questions to a considerable extent building upon the accomplishments of earlier ones"

(Sanders, 1974a). Thus, in at least some cases, literal questions might precede interpretive and applied questions. The literal questions would insure the basic literal understanding which is needed before interpretive and applied questions can be answered intelligently.

In general, reading specialists do not insist that all questions possess the characteristics described by Sanders. For example, the fact that a question lacks structure does not necessarily mark it as ineffective. Herber (1978) recommends that the use of more highly structured questions -- perhaps declarative statements which present possible responses from which students select -- should precede the use of less structured questions which allow students to determine their own responses. But it does seem reasonable to expect a balance of questions, both according to Sanders characteristics and according to Herber's levels of meaning.

Seminoff (1980) used Herber's (1978) three levels of meaning, and four of Sander's (1974a; 1974b) characteristics (clarity, objectivity, structure, and sequence) as the basis of her study of the questions in ten seventh grade social studies textbooks. She found that it was possible to train judges to reliably classify and describe textbook questions, based on the criteria of Herber and Sanders; and she found significant differences between at least some of the textbooks both in the

distribution of questions classified by level of meaning, and in the degree to which these questions possessed clarity, objectivity, structure, and sequence.

Method

The present study replicates Seminoff's (1980) study, but is based on eleventh grade American literature anthologies. A number of publishers' representatives and sales managers were asked to identify their major competitors in high school American Literature anthologies; from this informal survey, five anthologies were identified which met the following criteria:

(1) The selections were organized chronologically.

(2) The textbooks were hard-covered, major compiliations of American Literature.

(3) They were edited by one or more major editors.

(4) They contained questions in the student edition.

(5) They were recognized by their competitors as leaders in the field.

On closer study, one of the five texts was eliminated from the study. In both its teachers guide and its student edition, this anthology was vague about identifying which questions were intended to provide assistance to students in understanding and interpreting the selections; and it had a format which made it difficult to distinguish between questions intended for study or class discussion of the selection and questions intended for other purposes.

The selections in the remaining anthologies were classified

into one of three groups (genres): poetry; fiction; and non-fiction. Excerpts from novels were classified as fiction; and drama was excluded from consideration entirely, mainly because there were so few examples of it in the selected anthologies. Drama, of course, also presents special reading problems (related to its format and typography) which distinguish it from other types of fiction.

Four fiction and four non-fiction selections were randomly selected from each anthology. It was found that these selections represented about ten per-cent of each genre, so ten per-cent of the poems (from eight to fourteen per anthology) were also randomly selected. Although the number of poems was greater than the number of fiction and non-fiction selections, the fewer questions per poem (plus the shorter average length of the poems) was felt to justify this selection procedure.

The methodology from this point on paralleled that used by Seminoff (1980). All of the questions which accompanied each selection, and which were identified (either in the anthology itself or in the accompanying teacher's guide) as intended for study and/or discussion were selected. A form was created so that each selected question could be classified by levels of meaning (literal, interpretive, or applied), and rated on a five point scale for clarity, objectivity, structure, and sequence.

Four judges were selected to assist in the study. Each

judge had (1) a graduate degree, and (2) at least five years experience in secondary education. They were trained over a two month period to use the question classification and rating form. As part of the training, four test rounds were held using selections not part of the main study; the judges met regularly at one to two week intervals to discuss their responses in these test rounds, and to refine their thinking about the classifications and ratings.

Reliability coefficients were calculated for each of the four characteristics (clarity, objectivity, structure, and sequence) for each of the four test rounds. The coefficients were calculated using an ANOVA procedure described in Winer (1971, pp. 283-296). In the fourth round, three of the four coefficients were .90 or higher, while the fourth, for objectivity, was .72.

Reliability coefficients for the classifications — literal, interpretive, and applied -- were not calculated because the data was nominal; reliability in the main study was assured by including only questions on which at least three of the four judges agreed about the classification. Such majority agreement was reached on 89% of the questions.

In the main study, two rounds of data collection were held, with each judge working on half of the selections in each round. The order of presentation of the selections was varied between

judges to balance out fatigue effects. The judges first categorized each question according to level; then they rated each question for clarity, objectivity, and structure. Finally, they rated the complete set of questions for each selection for sequence. Each of the two rounds was completed over a two week period.

The data were compiled and the customary descriptive statistics calculated; these are presented in Tables One and Two for the levels of meanings, and in Tables Three and Four for the quality ratings. Finally, various statistical tests were calculated; these are reported and discussed in the next section of this study.

Findings

Tables Two through Five present simple descriptive statistics of the results. Tables Two and Three present the results for the ratings on each of the four characteristics; Tables Four and Five present the results for the classifications by level. On a scale of one (low) to five (high), the questions were rated an average of 1.86 on clarity, 3.11 on objectivity, 1.64 on structure, and 2.48 on sequence. 7.7% of the questions were classified literal, 86.3% interpretive, and 6.0% applied.

Table Six summarizes the results of a series of eight analysis of variance test of differences between texts and between genres on the four characteristics. Six of the eight tests were statistically significant at the .05 level; these included all four of the tests of differences between texts, and two of the four tests of differences between genres.

Where the analysis of variance tests were significant, a series of Newman-Keulls contrast tests were performed in an attempt to pin-point the specific differences which were significant. Because the Newman-Keulls test is a conservative test, and because the analysis of variance design was unbalanced, it was anticipated that some of these contrast tests might not be successful.

In these tests, there were significant differences (at the

.05 level) in clarity between texts two and three, and between texts two and four, with text two having lower ratings in each case. There were also significant differences in objectivity between text one and the other three texts, with text one having lower ratings in each case. There was in addition a significant difference in structure between text four and text one, with text four having the lower rating. And although the analysis of variance test was significant for sequence, the Newman-Keulls test did not pin-point significant differences between any specific texts.

The Newman-Keulls tests also revealed significant differences in objectivity between poetry and non-fiction, with poetry having a lower rating. Finally, the tests revealed significant differences in sequence between poetry on the one hand, and fiction and non-fiction on the other hand; poetry had the lower rating.

Conclusions and Recommendations

For both clarity and structure, the significant differences between texts are of little practical importance when the grand means for both characteristics are so low (1.86 and 1.64). Textbook questions in general are weak in both clarity and structure.

In addition, the texts were all weak in sequence; the highest text had a mean of only 2.73, well below the 3.00 midpoint of the rating scale. If the questions had been based on a well-worked out pedagogical approach, their sequence, and to a lesser extent their structure, should have been rated higher. The low ratings suggest that at best the pedagogical approach which guided the construction of these questions was poorly followed.

The preponderance of interpretive questions is also disturbing. It is true, of course, that a major purpose of most literature courses is interpretation. But can students be expected to interpret a literature selection before they have a literal level understanding of it? If not, wouldn't it be prudent to ask literal questions which help students understand before asking interpretive questions? And although "relevance" is not as fashionable as it once was, students might find literature more relevant if they spent some time responding to

questions on the applied level.

Perhaps the one non-negative finding in the study is the grand mean for objectivity (3.11). Literature is by nature subjective; a high level of objectivity may therefore not be necessary in literature questions. But on the other hand, by definition, questions which lack objectivity indicate the author's point of view; and when such questions come from a teacher or a textbook editor, they probably inhibit student expression of a subjective point of view.

As we expected, there were significant differences between both texts and genres on objectivity. One text had a significantiy lower rating on objectivity than the other three, and this fact might be important to a textbook selection committee.

Poetry had a lower rating on objectivity than the other two genres, and this difference might be appropriate since poetry is the least objective genre. But on the other hand, as we suggested in a previous paragraph, perhaps a higher level of objectivity in poetry questions would (paradoxically) encourage a higher level of subjectivity in student responses.

Possibly editors need to give more priority to the writing of textbook questions. In particular, perhaps they should be more careful about basing textbook questions on a specific pedagogical approach.

Textbook selection committees should consider these differences between instructional questions; they may even want to perform the kind of analysis done in this study as part of their selection process.

The situation described in this study is not encouraging. Questions in high school American literature anthologies are pedagogically inadequate in a number of ways. And we suspect that the anthologies reflect an inadequacy in the questions used in high school English classrooms. If asking the right question is 90% of learning, then English education may be operating at only 10% efficiency.

TABLE ONE

Descriptive Statistics by Genre
for the Four Characteristics of Questions

Qualities	Fiction	Non-Fiction	Poetry	Totals
Clarity:				
n	288	254	501	1043
mean	1.84	1.90	1.84	1.86
sd	0.73	0.70	0.72	0.72
Objectivity:				
n	288	254	501	1043
mean	3.07	3.41	2.99	3.11
sd	1.48	1.39	1.45	1.45
Structure:				
n	287	254	498	1039
mean	1.65	1.59	1.65	1.64
sd	0.57	0.54	0.60	0.57
Sequence:				
n	64	64	167	295
mean	3.02	2.48	2.27	2.48
sd	1.03	1.01	2.30	1.04

The qualities were rated on a five point scale ranging from one (low) to five (high).

TABLE TWO

Descriptive Statistics by Text
for the Four Characteristics of Questions

Qualities	Text One	Text Two	Text Three	Text Four	Totals
Clarity:					
n	283	276	212	272	1043
mean	1.87	1.71	1.95	1.92	1.86
sd	0.58	0.73	0.78	0.76	0.72
Objectivity:					
n	283	276	212	272	1043
mean	2.77	3.21	3.29	3.23	3.11
sd	1.47	1.47	1.39	1.41	1.45
Structure:					
n	283	275	212	269	1039
mean	1.73	1.60	1.68	1.55	1.64
sd	0.52	0.60	0.64	0.54	0.57
Sequence:					
n	83	63	67	82	295
mean	2.49	2.73	2.13	2.55	2.48
sd	1.06	2.00	0.89	1.04	1.04

The qualities were rated on a five point scale ranging from one (low) to five (high).

TABLE THREE

Frequency Distribution for "Level-of-Question" Categories According to Judges' Majority Classification by Genre and by Category

Categories	Fiction	Non-Fiction	Poetry	Totals
Literal:				
No. of Responses	4	8	6	18
Adj. Frequency	(5.7%)	(15.7%)	(5.4%)	(7.7%)
Interpretive:				
No. of Responses	64	39	98	201
Adj. Frequency	(91.4%)	(76.9%)	(87.5%)	(86.3%)
Applied:				
No. of Responses	2	4	8	14
Adj. Frequency	(2.9%)	(7.8%)	(7.1%)	(6.0%)
Totals:				
No. of Responses	70	51	112	233
Adj. Frequency	(100%)	(100%)	(100%)	(100%)

Only those questions were included in this tabulation on which a majority (three out of four) of the judges agreed on the classification.

TABLE FOUR

Frequency Distribution for "Level-of-Question" Categories According to Judges' Majority Classification by Text and by Category

Categories	Text One	Text Two	Text Three	Text Four	Totals
Literal:					
No. of Responses	3	9	5	1	18
Adj. Frequency	(4.9%)	(15.3%)	(10.4%)	(1.5%)	(7.7%)
Interpretive:					
No. of Responses	56	48	83	57	201
Adj. Frequency	(91.8%)	(81.4%)	(83.3%)	(87.7%)	(86.3%)
Applied:					
No. of Responses	2	2	3	7	14
Adj. Frequency	(3.3%)	(3.4%)	(6.3%)	(10.8%)	(6.0%)
Totals:					
No. of Responses	61	59	91	65	233
Adj. Frequency	(100%)	(100%)	(100%)	(100%)	(100%)

Only those questions were included in this tabulation on which a majority (three out of four) of the judges agreed on the classification.

TABLE FIVE

One Way ANOVA Tests
for the Differences in Qualities of Questions
Between Texts and Between Genres

Qualities	Texts F-Ratio	Texts Prob.	Genres F-Ratio	Genres Prob.
Clarity	5.912	.0005	0.562	.5700
Objectivity	7.296	.0001	7.586	.0005
Structure	5.432	.0010	0.934	.3935
Sequence	3.899	.0094	12.801	.0000

BIBLIOGRAPHY

Anderson, R.C., & Biddle, W.B. On asking people questions about what they are reading. In G. Bower (Ed.), Psychology of learning and motivation (Vol. 9). New York: Academic Press, 1975.

Bandman, B. What makes an educational question right or wrong. Philosophy of education 1967: proceedings of the twenty-third annual meeting of the Philosophy of Education Society, 1967, 127-140.

Belleck, A.A.; Kliebard, H.; Hyman, R.; & Smith, F.L. The language of the classroom. New York: Teachers College Press, 1966.

Browne, M. Neil, & Keeley, Stuart. Asking the right questions. Englewood Cliffs, NJ: Prentice-Hall, 1981.

Henry, G.H. Teaching reading as concept development: emphasis on affective thinking. Newark, DE: International Reading Association, 1974.

Herber, H.L., & Nelson, J.B. Questioning is not the answer. Journal of Reading, 1975, 18, 512-517.

Herber, H.L. Teaching reading in content areas (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, 1978.

Hill, W.R. Secondard school reading: process, program, procedure. Boston, MS: Allyn & Bacon, 1979.

Hoether, J. Teacher questioning behavior in nine junior high school English classes. Research in the Teaching of English, 1968, 2, 99-108.

Hunkins, F.P. Questioning strategies and techniques. Boston, MS: Allyn & Bacon, 1972.

Hyman, R.T. Ways of teaching. New York: Lippincott, 1970.

Riegle, R.P. The logical characteristics of classroom questions. Philosophy of education 1975: proceedings of the thirty-first annual meeting of the Philosophy of Education Society. 1975, 101-112.

Sanders, P.L. When teachers talk and students (sometimes)

listen. English Journal, 1974, 63, 80-81 (a).

Sanders, P.L. Who's responsible for what? one definition of the English teacher's role. English Journal, 1974, 63, 94-95 (b).

Santa, Carol M., & Hayes, Bernard L., (eds.). Children's prose comprehension: research and practice. Newark, DE: International Reading Association, 1981.

Squire, J.R., & Applebee, R. High school English instruction today. New York: Appleton-Century-Crofts, 1968.

Wease, H. Questioning the genesis of teaching and learning. High School Journal, 1976, 59, 258-265.

Winer, B.J. Statistical Principles in Experimental Design. Second Edition. New York: McGraw-Hill, 1971.

Winne, P.H. Experiments relating teacher's use of higher cognitive questions to ability achievement. Review of Educational Research, 1979, 49, 13-50.